Some Footnotes on Protein Synthesis

A Note for the RNA Tie Club

- by -

F.H.C. Crick and S. Brenner

M.R.C. Unit for Molecular Biology,
Cavendish Laboratory,
Cambridge,
England.

December, 1959.

Summary

In this note we point out:

(1) that soluble RNA may have a DNA-like secondary structure in order to cover up the "adaptor sequence" when no amino acid is attached to the RNA.

(2) that at some stage the amino acid may be attached to the N_1 of the pseudo-uridine of soluble RNA.

(3) that two particles of different size and shape can be constructed from identical protein subunits if combined with RNA molecules of different length. In particular both the 50 S and the 30 S ribosomal particles may be incomplete parts of a spherical shell.

The object of this brief note is to add a few details to some of our ideas on protein synthesis.

Soluble RNA

The first idea that needs some revision is the "adaptor hypothesis". This originally sprang from the difficulty of conceiving an RNA template with 20 specific cavities, one for each amino acid. It was suggested that 20 specific "adaptor molecules" would be required to position the amino acids, and that each amino acid would be joined to its own special adaptor by a special enzyme. The nature of these adaptors was initially left open, but it was later pointed out that the obvious choice was a small polynucleotide which would combine, by base-pairing, with the proper parts of the base sequence of the "template" RNA.

The subsequent discovery of the 20 specific activating enzymes, and their role in transferring the amino acids specifically to molecules of soluble RNA at first supported these ideas. Further work, however, has uncovered some difficulties. They are:

(1) The amino acids become attached to the terminal ribose of the soluble RNA, yet the sequence of bases at that end appears to be the same for all (or almost all) molecules of soluble RNA, namely Adenine Cytosine (Cytosine). One might have expected that this part of the sequence was specific for each amino acid.

(2) The soluble RNA has a higher molecular weight than was expected. Estimates vary from 50 to 100 bases, with 80 to 90 as the more likely values. One might have expected any number greater than 2 but less than (say) 10.

We can think of reasons for explaining the need for the common terminal sequence - to fit on to a standard enzyme, to provide a link to bridge a distance, etc., - but as they are all rather obvious and as there is nothing much to choose between them we shall not discuss them further here.

On reflecting on the problem of the apparently excessive size of the soluble RNA, however, it became apparent that a rather important requirement had been omitted. It would

delay protein synthesis unduly if an empty adaptor (i.e. without its amino acid) could fit onto the template, and there was nothing in the original theory to prevent this.

This difficulty can be got round in a number of ways. The obvious thing to do is to cover up the specific base-sequence, needed for base pairing with the template, until the amino acid is attached. One rather neat way of doing this suggests itself. This is to make the soluble RNA fold back on itself to form a DNA-like structure. (We know from studies of "Poly A plus Poly U" that a polynucleotide with a ribose backbone can take up a configuration like DNA). This requirement could well account for the length of the RNA being greater than expected. The idea that soluble RNA may have some secondary structure is not new. Dr. Paul Berg thought at one time that he had some evidence for it (and so, more recently, have others) and it had been suggested to us in discussion by Dr. Boman and by Dr. Bach. The only novelty in our idea is that it suggests a reason for such a secondary structure, if it exists.

The idea, then, is that without the amino acid the molecules of soluble RNA fold on themselves so that their "Adaptor sequence" is covered. At some time after the amino acid is attached this sequence is uncovered, and can then base-paid with the template RNA of the ribosomes.

Is there any evidence that soluble RNA has secondary structure? This is too recent and complex a subject to discuss here, but we may note that the base composition of soluble RNA, determined by three laboratories (Kendrick Smith, David Dunn and Jim Offengand, with their various collaborators - personal communications) have shown that the base ratios tend towards the DNA rule; that is

a $\simeq$ U plus pseudo-U and G $\simeq$ C.

(Incidentally if soluble RNA resembles DNA it may act as a template for its own replication).

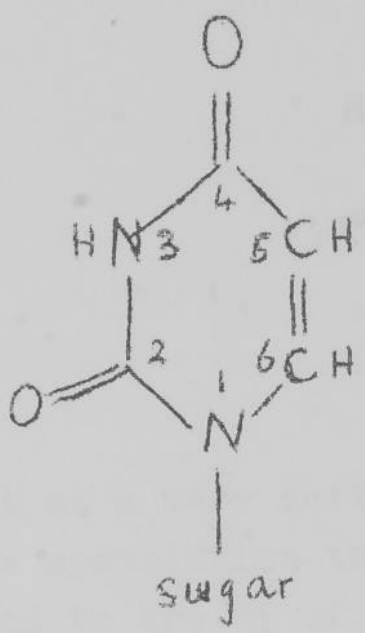
There is one difficulty about this idea which is imaginary. This is: how does the activating enzyme recognise the correct molecules of S-RNA if the adaptor

sequence is covered? The answer is that a protein can in principle recognise a pair of bases, by making contact in the deep groove of the DNA-like structure, and so can distinguish one double-helix from another, with a different base sequence, without necessarily unpairing the bases. There does not seem to be any simple way in which a polynucleotide chain can do this.

Before we discuss this further let us consider another peculiarity of the soluble RNA.

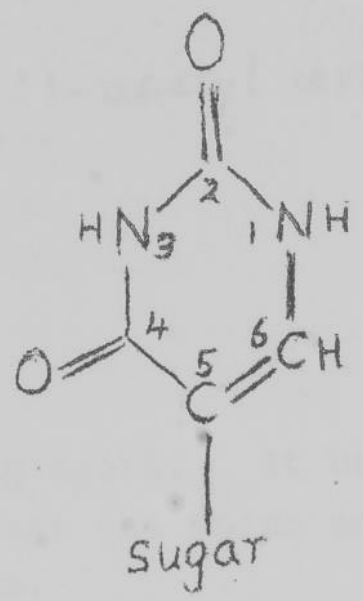
This is the surprislingly large amount of pseudo-uridine; that is, uracil attached to the C_1' of the ribose by the C_5 rather than by the N_1, thus making a C-C bond between sugar and base. Some 3% of the bases in soluble RNA are like this. Moreover there is none of it in the RNA of TMV, and little (probably none) of it in pure ribosomal RNA.

It is instructive to draw the formula of this attachment (Fig. 1a) and compare it with the way in which uracil is attached.



uridine

Fig 1 (a)



pseudouridine.

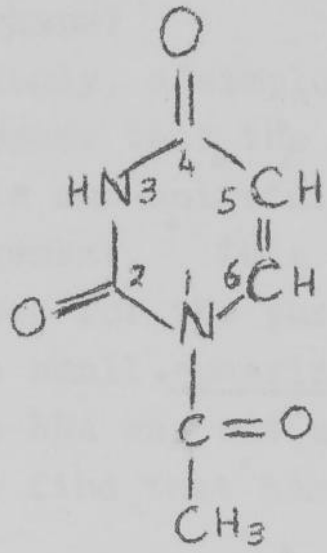
Fig 1 (b)

One is surprised to find that due to the symmetry of uracil it is as if one had kept all the atoms in fixed places but had exchanged N_1 and C_5. Note that this puts an active atom - a nitrogen - in the old position of C_5. Also note that the part of the molecule which is used to form a base-pair with adenine is unaffected by this change.

It is impossible to believe that this is an accident or is done for trivial reasons. A little thymine may slip into RNA, or a methyl group may be tacked on here and there, and nobody need get excited. But to forge a carbon-carbon bond, and to provide an appreciable amount in one special type of RNA ..... one does not have to be a thereotician to hear Nature when she shouts at one.

Moreover it will not have escaped the notice of those of you who manage to read the JBC from cover to cover that it has been shown (Spector and Keller, J.B.C. 232, 185 (1958) that the molecule

Fig 2.



1-acetyl uracil

can act as a very active acetylating agent. It becomes an obvious speculation that at some stage the amino acid is attached to the N_1 of pseudo-uridine.

Whether these two ideas can usefully be united is another matter. One could speculate that it is the transfer of the amino acid to the pseudo-uridine (requiring GTP ?) which disturbs the secondary structure and uncovers the adaptor, but as we do not know whether the entire molecule of soluble RNA goes into the ribosomes or only part of it a large number of possibilities exist.

Ribosomes.

Our original speculations about these particles was that the protein part was mainly "structural", and would be made of sub-units, and that the particles might well have cubic symmetry, or some approximation to it.

It was therefore a surprise to find that the 70 S particle breaks up into a 50 S and a 30 S particle. Why should there be two unequal particles, and how are they related?

Nobody seems to have produced a very good reason for the two parts of the 70 S particle, except the idea that the 50 S is a box and the 30 S is its lid, and that the two come apart momentarily, from time to time, to let finished protein molecules escape.

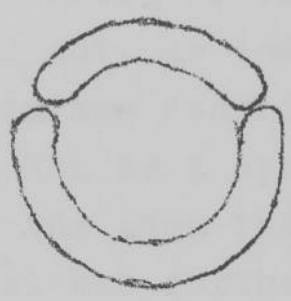
As to the manner of construction it would not be surprising if there was one protein for the 50 S particle and one for the 30 S, but recent work (The Harvard Biology group - personal communication) suggests that the protein of these two particles are rather similar. How then do they come to have such a different size and shape?

Rather surprisingly, a simple answer is possible. It is only necessary to assume that the particle cannot easily be produced from protein sub-units alone, but needs RNA to stabilise the arrangement. This is in any case true for the rods of TMV. Suppose for the purposes of explanation, that it were true for the small spherical RNA viruses. Then if, before assembly, the RNA was cut exactly in half one might reasonably expect to find that hemispheres would be produced instead of spheres.

On this view, then, each of the two portions of a ribosome consists of part of a spherical shell; the size of each would be determined by the size of the RNA, so that the protein/RNA ratio would be the same for both parts, as it found to be the case. It is not essential for the protein sub-units of the two parts to be identical, but one would expect them to be similar.

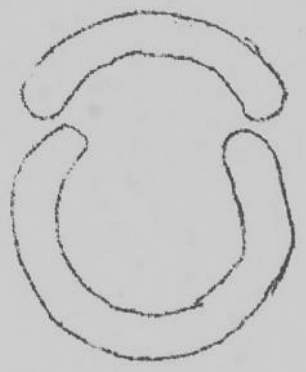
At first sight the structure might be

Fig. 3

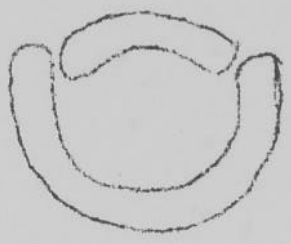


but this makes it difficult to see why the two parts should not be equal. Other alternatives are :

Fig. (a)



(b)



These have the advantage that the contact between the two parts is different from that between the sub-units of each part. Curiously enough recent electron micrographs (Hugh Huxley - personal communication) do not look totally unlike Fig. 4(a). It is in any case known that the smaller particles (30 S) is more asymmetrical than the bigger one (50 S).

This sort of idea would explain very well why the molecular weight of the RNA has, apparently, two fixed sizes (leaving aside the difficult question of the RNA breaking down into further sub-units) and not, as one might expect, a continuous range of sizes. It raises considerable difficulties when one comes to consider the RNA as a specific template. One is naturally inclined to the idea that only <u>part</u> of the ribosomal RNA acts as a template, but further than that we shall not venture at the moment.